What is claimed is:

1.

A method of elucidating a protein expression profile of a test cell comprising:

introducing to the genome of said test cell a polynucleotide construct, said polynucleotide construct encoding an assay marker peptide, the expression of which is obtained only upon integration of said polynucleotide construct to an actively transcribing genome region of said test cell;

sorting said cell based upon the expression of said assay marker peptide;

correlating said expression with a genetic locus and

comparing said expression with a reference cell to identify a protein expression profile for the test cell indicative of the test cell.

2.

The method of claim 1 wherein said polynucleotide construct is a vector.

3.

The method of claim 1 wherein said vector is a viral vector.

4.

The method of claim 3 wherein said viral vector is based upon a virus selected from the group consisting of: retrovirus, lenti-virus, adenovirus or adeno-associated virus.

5.

The vector of claim 4 wherein said vector is a retroviral vector.

6.

The method of claim 1 wherein said assay marker peptide is a fluorescent protein.

7.

The marker of claim 6 wherein said peptide is green fluorescent protein.

8.

The marker of claim 7 wherein said marker is a hrGFP.

9.

The method of claim 1 wherein said polynucleotide construct further comprises a polyadenylation signal.

10.

The method of claim 1 wherein said polynucleotide construct further comprises an internal ribosome entry site.

11.

The method of claim 1 wherein said polynucleotide construct further comprises a splice acceptor site.

12.

The method of claim 1 wherein said polynucleotide construct further comprises a selectable marker gene.

13.

The method of claim 1 wherein said sorting is accomplished by separating said cells on the basis of chemiluminescence, fluorescence, or mechanical means.

14.

The method of claim 13 wherein said sorting is by fluorescence activated flow cytometry.

15.

The method of claim 1 wherein said test cell is a cancer cell and said reference cell is a normal cell.

16.

The method of claim 15 wherein said test cell is a colon cell and said reference cell is a colon cancer cell.

17.

The method of claim 1 wherein said correlating step comprises the following steps:

isolating nucleic acid material from said separated cell;

cleaving said nucleic acid material so that the region of integration of said polynucleotide is separated;

amplifying said cleaved region to form an amplicon comprising known and unknown sequence;

sequencing said amplicon; and

correlating said unknown amplicon sequence with a sequence database to identify a genetic locus associated with said amplicon.

18.

The method of claim 17 wherein said step of cleaving said nucleic acid includes cleaving said nucleic acid so that the inserted polynucleotide construct is cleaved once and flanking nucleic acid sequence is cleaved in a region contiguous to the inserted polynucleotide.

19.

The method of claim 18 wherein said step of cleaving said nucleic acid includes cleaving said nucleic acid at a site that is a fixed distance from said integrated polynucleotide.

20.

The method of claim 19 further comprising the step of: ligating the ends of said cleaved nucleic acid so that a concatamer is formed.

21.

A method of elucidating a relative quantitative protein expression profile of a population of test cells comprising:

introducing to the genome of a plurality of said test cells a polynucleotide construct, said polynucleotide construct encoding an assay marker peptide, the expression of which is obtained only upon integration of said polynucleotide construct to an actively transcribing genome region of said test cell population;

sorting said cells based upon the expression level of said assay marker peptide;

correlating said expression with a genetic locus and

comparing said expression with a reference cell population to identify a protein expression profile for the test cell population indicative of the test cell population.

22.

The method of claim 1 wherein said polynucleotide construct is a vector.

23.

The method of claim 22 wherein said vector is a viral vector.

24.

The method of claim 23 wherein said viral vector is based upon a virus selected from the group consisting of: retrovirus, lenti-virus, adenovirus or adeno-associated virus.

25.

The vector of claim 24 wherein said vector is a retroviral vector.

26.

The method of claim 21 wherein said assay marker peptide is a fluorescent protein.

27.

The marker of claim 26 wherein said peptide is green fluorescent protein.

28.

The marker of claim 27 wherein said marker is a hrGFP.

29.

The method of claim 21 wherein said polynucleotide construct further comprises a polyadenylation signal.

30.

The method of claim 21 wherein said polynucleotide construct further comprises an internal ribosome entry site.

31.

The method of claim 21 wherein said polynucleotide construct further comprises a splice acceptor site.

32.

The method of claim 21 wherein said polynucleotide construct further comprises a selectable marker gene.

33.

The method of claim 21 wherein said sorting is accomplished by separating said cells on the basis of chemiluminescence, fluorescence, or mechanical means.

34.

The method of claim 33 wherein said sorting is by fluorescence activated flow cytometry.

35.

The method of claim 21 wherein said test cell population is a cancer cell population and said reference cell population is a normal cell population.

36.

The method of claim 35 wherein said test cell population is a colon cell and said reference cell population is a colon cancer cell.

37.

The method of claim 21 wherein said correlating step comprises the following steps:

isolating nucleic acid material from said separated cells;

cleaving said nucleic acid material so that the region of integration of said polynucleotide is separated;

amplifying said cleaved region to form an amplicon comprising known and unknown sequence;

sequencing said amplicon; and

correlating said unknown amplicon sequence with a sequence database to identify a genetic locus associated with said amplicon.

38.

The method of claim 37 wherein said step of cleaving said nucleic acid includes cleaving said nucleic acid so that the inserted polynucleotide construct is cleaved once and flanking nucleic acid sequence is cleaved in a region contiguous to the inserted polynucleotide.

39.

The method of claim 38 wherein said step of cleaving said nucleic acid includes cleaving said nucleic acid at a site that is a fixed distance from said integrated polynucleotide.

40.

The method of claim 39 further comprising the step of: ligating the ends of said cleaved nucleic acid so that a concatamer is formed.

41.

The method of claim 1 wherein said correlating step comprises the following steps:

isolating nucleic acid material from said separated cell;

cleaving said nucleic acid material so that the region of integration of said polynucleotide is separated;

amplifying said cleaved region to form an amplicon comprising known and unknown sequence;

sequencing said amplicon; and

correlating said unknown amplicon sequence with a sequence database to identify a genetic locus associated with said amplicon.

42.

The method of claim 41 wherein said step of cleaving said nucleic acid includes cleaving said nucleic acid so that the inserted polynucleotide construct is cleaved once and flanking nucleic acid sequence is cleaved in a region contiguous to the inserted polynucleotide.

43.

The method of claim 42 wherein said step of cleaving said nucleic acid includes cleaving said nucleic acid at a site that is a fixed distance from said integrated polynucleotide.

44.

The method of claim 43 further comprising the step of: ligating the ends of said cleaved nucleic acid so that a concatamer is formed.

45.

A polynucleotide construct pGT5A.

46.

A polynucleotide construct pGT5AH.

47.

A polynucleotide construct pGT5Z.

48.

A polynucleotide construct pGT7A.

49.

A polynucleotide construct pGT7AH.

50.

A polynucleotide construct pGT7Z.

51.

A method of elucidating a protein expression profile of a test cell comprising:

introducing to the genome of said test cell a polynucleotide construct, said polynucleotide construct encoding an assay marker peptide, the expression of which is obtained only upon integration of said polynucleotide construct to an actively transcribing genome region of said test cell;

sorting said cell based upon the expression of said assay marker peptide;

sequencing the region marked by said expression marker using sequence tag acquisition and reporting method;

correlating said expression with a genetic locus and

comparing said sequence with a reference cell to identify a protein expression profile for the test cell indicative of the test cell.

52.

A method of elucidating a protein expression profile of a test cell comprising:

introducing to the genome of said test cell a polynucleotide construct, said polynucleotide construct encoding an assay marker peptide, the expression of which is obtained only upon integration of said polynucleotide construct to an actively transcribing genome region of said test cell;

sorting said cell based upon the expression of said assay marker peptide;

sequencing the region marked by said expression marker using serial analysis of viral integration method;

correlating said sequence with a genetic locus and

comparing said expression with a reference cell to identify a protein expression profile for the test cell indicative of the test cell.